Mission Shares Policy on Pleiades

Category: PBS on Pleiades

DRAFT

This article is being reviewed for completeness and technical accuracy.

Mission Directorate shares have been implemented on Pleiades since Feb. 10, 2009. Implementing shares guarantees that each Mission Directorate gets its fair share of resources.

The share to which a job is assigned is based on the GID used by the job. Once all the cores within a Mission Directorate's share have been assigned, other jobs assigned to that share must wait, even if cores are available in a different Mission Directorate's share, with the following exception:

When a Mission Directorate is not using all of its cores, other Mission Directorates can borrow those cores, but only for jobs that will finish within 4 hours. When part of the resource is unavailable, the total number of cores decreases, and each Mission Directorate loses a proportionate number of cores.

You can display the share distribution by adding the "-W shares=-" option to the qstat command. For example:

%qstat -W shares=-

Group	Share%	Use%	Share	Exempt	Use	Avail	Borrowed	Ratio	Waiting
Overall	100	0	159748	0	960	158788	0	0.01	960
ARMD	24	18	38109	0	29680	8429	0	0.78	22512
HEOMD	23	21	36521	0	34312	2209	0	0.94	28416
SMD	51	50	80981	0	80968	13	0	1.00	113920
NAS	2	0	3175	0	0	3175	0	0.00	20240

Mission shares are calculated by combining the mission's HECC share of the shared assets combined with the mission-specific assets. The mission shares on Oct 3, 2011 are shown in the second column of the above display. The amount of resources used and borrowed by each mission and resources each mission is waiting for are also displayed.

An in-house utility, *qs*, provide similar information with details that break the resources into the Harpertown, Nehalem-EP and Westmere-EP processor types and is available at /u/scicon/tools/bin/qs.

The -h option of qs provides instructions on how to use it:

```
% /u/scicon/tools/bin/qs -h
usage: qs [-u] [-n N] [-b] [-p] [-d] [-r] [-f M,N] [-q N] [-t] [-v] [-h] [--file f]
                   : show used resources only; don't show queued jobs
                   : show time remaining before N nodes are free
          -n N
                   : order segments in bars to help understand borrowing
          -b
                   : plain output: i.e. no colors or highlights
          -р
                   : darker colored resource bars (for a light background)
                  : use Reverse video for displaying resource bars
          -r
                  : highlight nodes for jobs that finish in <= M minutes
          -f M, N
                     and <= N minutes [default M=60, N=240]
                     (0 turns off highlighting)
                   : highlight nodes for jobs queued in last N minutes [3]
          -q N
                     (0 turns off highlighting)
          -t
                   : show time remaining & nodes used for each running job
          --file f : reserved for debugging
                   : (verbose) provide explanation of display elements
                   : provide this message
```

Here is a sample output file of *qs*:

```
%qs
               100% of mission share
ARMD
                           ыныныныыныныныныныны
r: 77%
      <u> h</u>hhhhhրաաաաաաաաաաաաաա
                           w: 59%
      hhhhhhnwwwwwwwwwwwww
                           имимимимымымымымымы
имимимымымымымымы
      hhhhhhnwwwwwwwwwwwww
      hhhhhnnաաաաաաաաաաաաա
                            hhhhhnաասաաաաաաաաաա
                           ЫНЫНЫЫЫЫЫЫЫЫЫЫЫЫЫ
      hhhhhnրաաաատատատատատա
HEOMD.
r: 93%
ω: 77%
      <u>հի</u>հիհիհիհիդաաաաաաաաաաաաաա
                               НННимимимимимимимимимимимим
      hhhhhhhhhhhhաաաաաաաաաաաաաա
                               <u>ի</u>հիհիհիհիհիհաաաաաաաաաաաաաա
      <u>ի</u>իհիհիիիիի<mark>ցաս</mark>աաաաաաաաաաաա
<u>ի</u>հիհիհիհիո<mark>սա</mark>աաաաաաաաաաա
                               SMD
      r: 99%
w:141%
      <u>h</u>hhhhhhhnnnnnաաաաաաաաաաաաաա
                                НННННNNNNNNAAAAAAAAAAAA
                                HHHHHNNNNNAAAAAAA
      <u>h</u>hhhhhhnnnnnwwwwwwwwwwwwww
      hhhhhhhnnnnnшшшышшшшшшшшшш
                                hhhhhhnnnnnwwwwwwwwwwwwww
      hhhhhhhnnnnwwwwwwwwwwwwwww
                                 hhhhhhnnnnuuuuuuuuuuuuuuu
                                 hhhhhhhnnnnwwwwwwwwwwwwwww
                                 -waiting-
Total Nodes Used ..... har:5212
                                        neh:1051
                                                wes:3636
Unused Nodes ...... har: 288(+63 bigmem) neh: 227
                                                wes: 243(+62 gpu)
                                             0
Unused, but limited by Resv.. har:
                                        neh:
                                               wes:
                                                     -0
Unused in <u>devel</u> queue ...... har։
hhhhnnnnnnnաատաատատա<u>տատատա</u>
                                        neh:
                                             0
                                               wes: 165
      hhhhnnnnnwwwwwwwww<u>wwwwww</u>
      hhhhnnnnnաաաաաաաաաաաաաա
      Each letter (h,n,w,H,N,W) ~= the same number of SBUs/hr ==> in nodes: h/H ~= 24.3 Har.; n/N ~= 12.2 Neh.; w/W ~= 8.1 Wes.
```

Article ID: 168

Last updated: 03 Oct, 2011

Computing at NAS -> Running Jobs with PBS -> PBS on Pleiades -> Mission Shares Policy

on Pleiades

http://www.nas.nasa.gov/hecc/support/kb/entry/168/?ajax=1